Singular Landscapes: in honor of Bernard Teissier
22-26 June, 2015

# Geometry of some functional architectures of vision

Jean Petitot
CAMS, EHESS, Paris

## Bernard and visual neuroscience

Bernard helped greatly the developement of geometrical models in visual neuroscience.

- In 1991 he organized the first seminars on these topics at the ENS and founded in 1999 with Giuseppe Longo the seminar *Geometry and Cognition*.

- From 1993 on, he organized at the *Treilles Foundation* many workshops with specialists such as Jean-Michel Morel, David Mumford, Gérard Toulouse, Stéphane Mallat, Yves Frégnac, Jean Lorenceau, Olivier Faugeras.

- He organised also in 1998 with J.-M. Morel and D. Mumford a special quarter *Mathematical Questions on Signal and Image processing* at the IHP.

- He worked with Alain Berthoz at the College de France (Daniel Bennequin worked also a lot there on geometrical models in visual neuroscience).

In this talk I would try to explain some aspects of Neurogeometry, concerning the link between natural low level vision of mammals and geometrical concepts such as fibrations, singularities, contact structure, polarized Heisenberg group, sub-Riemannian geometry, noncommutative harmonic analysis, etc.
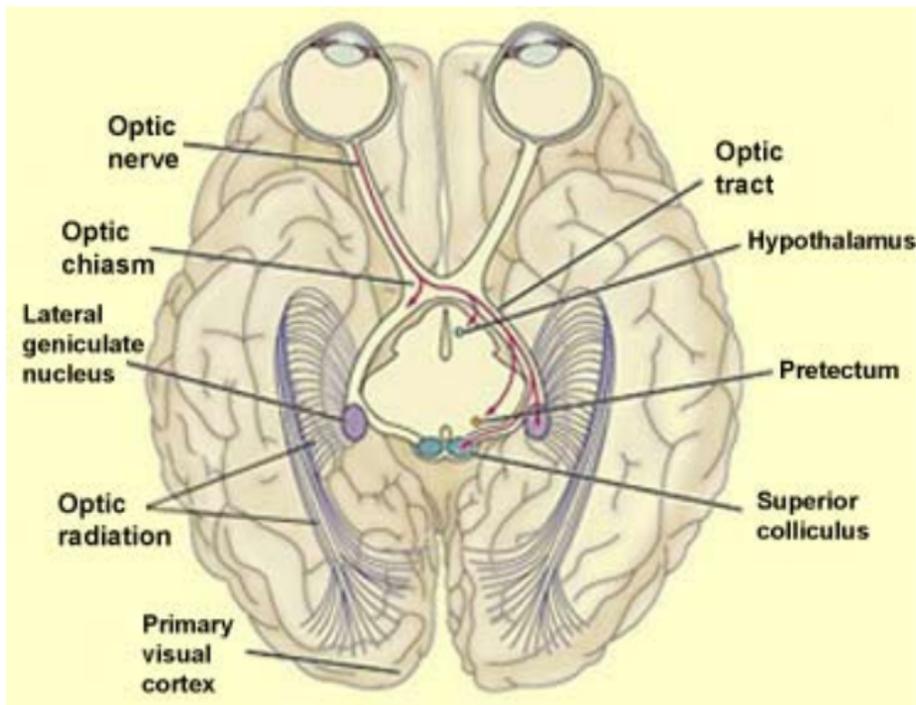
I will introduce some very basic and elementary experimental facts and theoretical concepts.

QUESTION:
How the visual brain can be a neural geometric engine?

## The visual brain

Here is an image of the human brain. It shows the neural pathways from the retina to the lateral geniculate nucleus (thalamic relay) and then to the occipital primary visual cortex (area $V1$).

# fMRI of human V1

fMRI of the retinotopic projection of a visual hemifield on the corresponding V1 (human) hemisphere. Concentric circles and rays are coded by colors.



20 deg
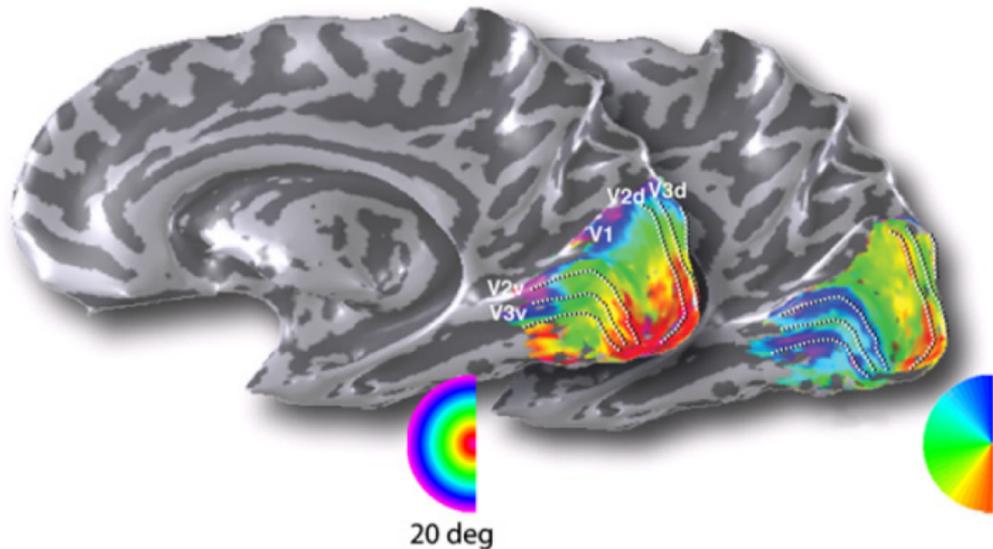
## Area V1

I will focus on the first primary area V1 (cat's area 17).

I will say nothing of

1. the retino-geniculo-cortical pathway projecting retinotopically the retina onto V1 (it is a conformal map);
2. the post-V1 processing of visual images by other cortical areas.

This restriction is of course a drastic limitation and simplification: reality is much more complex. But it can be justified by Mumford's "high-resolution buffer hypothesis" according to which V1 is much more than a simple "bottom-up early-module", and is essential to any visual processing requiring a fine resolution.

The geometry of visual perception provides a lot of evidences for some sort of neural implementation of differential and integration routines (detection of tangents, detection of curvature, integration of curves, etc.).

But what type of implementation?

Visual cortical neurons are very local detectors and even "point processors" (Jan Koenderink). They are only able to code a single numerical value by means of their firing rate.

They CANNOT implement differential routines.

But they are connected in a very *specific* way, called a *functional architecture*, and compose very complex neural nets.

Therefore, we must understand how such nets of neurons can detect very local differential features and integrate them into the global geometric structures of perception?

The classical intuitions of "differentiation" and "integration" cannot be used. We need much elaborate concepts of differential geometry.

The crucial discovery of a functional architecture of V1 was made in 1959 by David Hubel and Torsten Wiesel (1981 Nobel Prizes).

Recording the activity of cells of V1 (cat's area 17) they discovered, almost by accident, that some of these cells (they called "simple" ) were activated by a bar of a well defined orientation.

Here are some seconds of this crucial experiment

HW1

Then they discovered that the selectivity to orientation is like a
Gaussian with a peak of cell activity for the preferred orientation
and no cell activity for the orthogonal orientation.

HW2

Hubel and Wiesel analyzed the dependence of the orientation tuning curve w.r.t. the position of the bar, its orientation, its thickness, its length, its contrast, etc.

They delimitated the position and the size of the receptive field of a simple cell, that is the small region of the visual field to which it is correlated.

In a very rough linear approximation, neurons act as filters on the optical signal transduced by the photoreceptors of the retina. Their *receptive profile* is their transfer function.

They are well modeled by Gabor patches or derivatives of Gaussian.

# Example of receptive field

Here is an example. Left: Recording of level sets (Gregory DeAngelis, Berkeley). Right: model (third derivative $\varphi(x, y) = \frac{\partial^3 G}{\partial x^3}$).
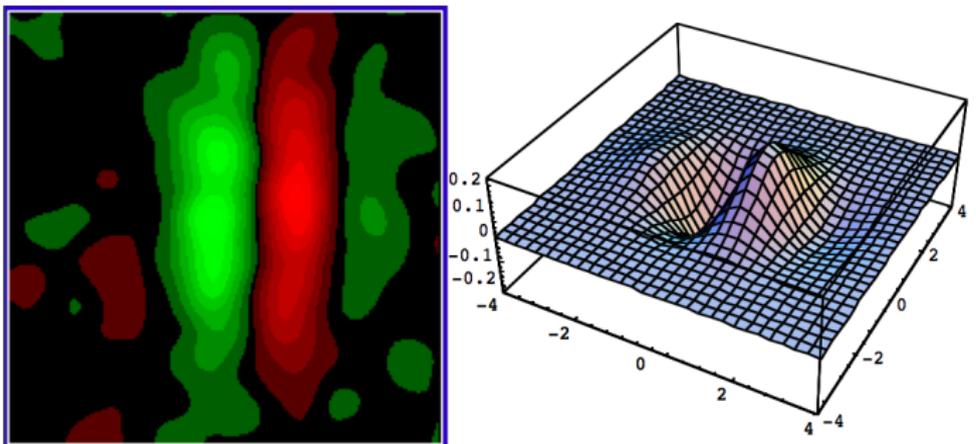
The filtering of the signal is like a *wavelet analysis* using oriented wavelets.

Simple cells of V1 are parametrized by triples $(a, p)$ where $a = (x, y)$ is a position in the retina or the visual field (that one can identify to $\mathbb{R}^2$) and $p$ is an orientation at $a$.
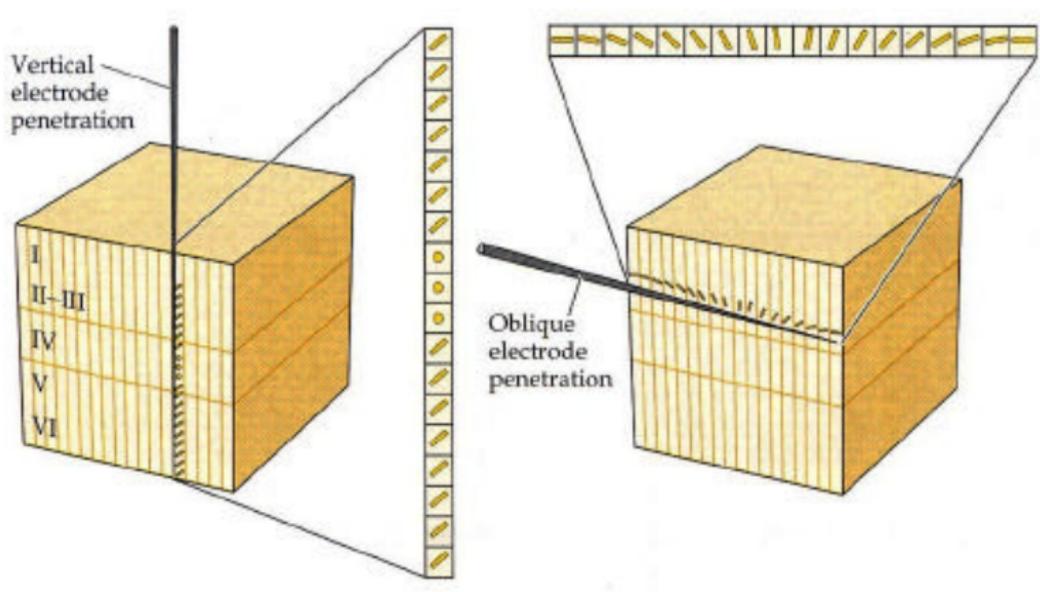
So, simple cells of V1 constitute a *field of orientations*.

This structure is the basis for the "*functional architecture*" of V1.

Immediately the question arises: what is the structure of this field?

# Hubel & Wiesel crucial experiments 2

The second breakthrough made by Hubel and Wiesel was that orientations vary as continuously as possible: along a "horizontal" penetration, position remains *constant* while orientation *rotates* regularly until it meets some singularity.

It was a great discovery.

Hence the first idealization of the functional architecture:

neurons detecting all the orientations $p$ at the same position $a$ of $\mathbb{R}^2$ constitute an anatomically well defined small neural module called an "orientation hypercolumn".

## Geometrical interpretation

This means that the *fiber bundle* $\pi : \mathbb{V} = \mathbb{R}^2 \times P \to \mathbb{R}^2$ (where the fiber $P$ is the projective line of orientations) is *neurally* implemented.

This is mathematically trivial but not neurophysiologically trivial as a result of evolution.

# Braitenberg reconstruction

One of the first plausible reconstruction of a concrete orientation field from the sparse data provided by a network of electrodes was *infered* in 1979 by Valentino Braitenberg.

Taking into account the fact that along a linear tangential penetration the chirality of rotating orientations could change, Braitenberg reached the conclusion:

> "We believe that the most natural explanation of the facts observed would be in terms of orientations arranged with circular symmetry around centers, either radially or along concentric circles."

## Orientation centers

The introduction of centers of orientation explains the inversion of chirality along a linear penetration.
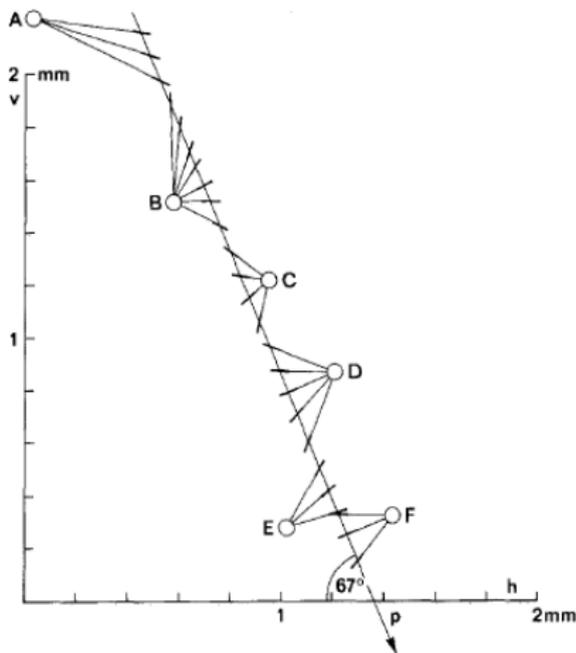
Braitenberg's inference has been strikingly confirmed in the 1990s by the revolution of brain imagery (fMRI).

Here is the functional architecture of the area $V1$ of a tree-shrew (tupaya) obtained by "in vivo optical imaging" (William Bosking).

B

Lateral

Rostral

1 mm

C

200 μm
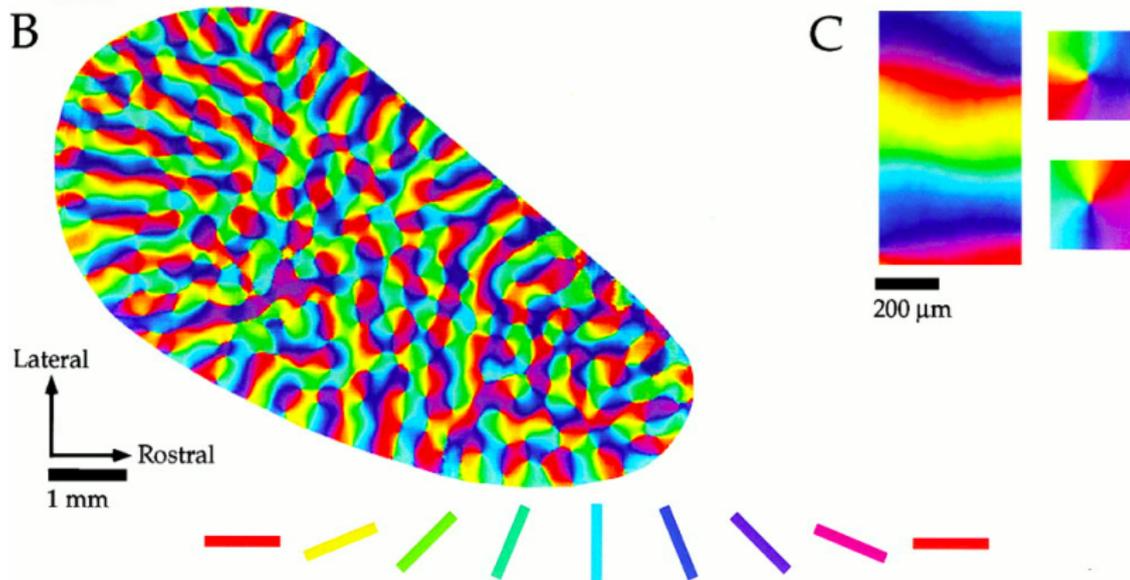
# Pinwheels figure

- The plane is $V1$,
- A colored point represents the mean of a small group of real neurons (mesoscale).
- Colors code for the preferred orientation at each point.
- The field of isochromatic lines (i.e. iso-orientation lines) is organized by a lattice of *singular points* called *pinwheels* where all orientations meet.
- There exist a "mesh" of the lattice of pinwheels (a sort of characteristic length).
- Pinwheels have a chirality.
- Adjacent pinwheels have opposed chirality.

If $\theta$ is the angle of a pinwheel ray, the associated orientation varies, up to a constant, as $\pm\theta/2$.

In the following picture due to Shmuel (cats area 17), orientations are coded by colors but are also represented by white segments.

We observe very well the two types of generic singularities of 1D foliations in the plane: end points ($+\theta/2$) and triple points ($-\theta/2$). They correspond to the two possible chiralities of pinwheels.

# Shmuel figure
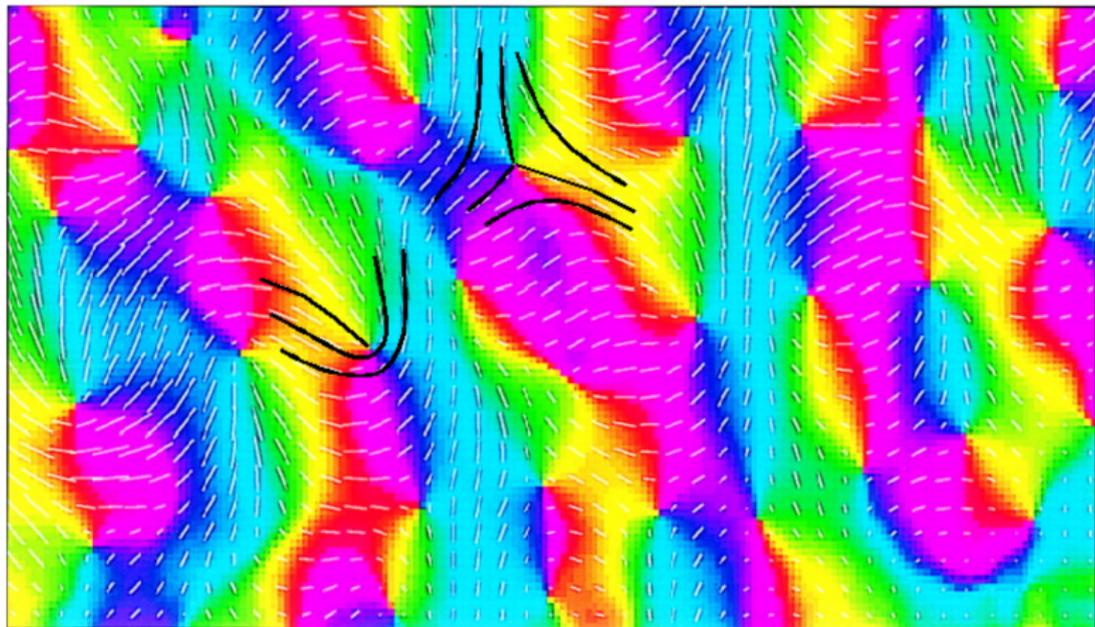
## Wolf-Geisel models

There exist beautiful models of pinwheels. They are analogous to
dislocations of phase fields in optics (see Mike Berry's works).
Fred Wolf and Theo Geisel proposed a model using a complex field

$$Z : \mathbb{R}^2 \to \mathbb{C}, \; a = \rho e^{i\theta} \mapsto r(a) e^{i\varphi(a)}$$

where the spatial phase $\varphi(a)$ codes the orientation ($\varphi(a)$ varies as
$\pm\theta/2$ near singular points) and the module $r(a)$ codes the
orientation selectivity.

Under the (non trivial) hypothesis that there is no selectivity at the
pinwheel singularities, these singular points are zeroes of the field.

If $Z = X + iY$, they are the intersections of the curves $X = 0$ and
$Y = 0$.

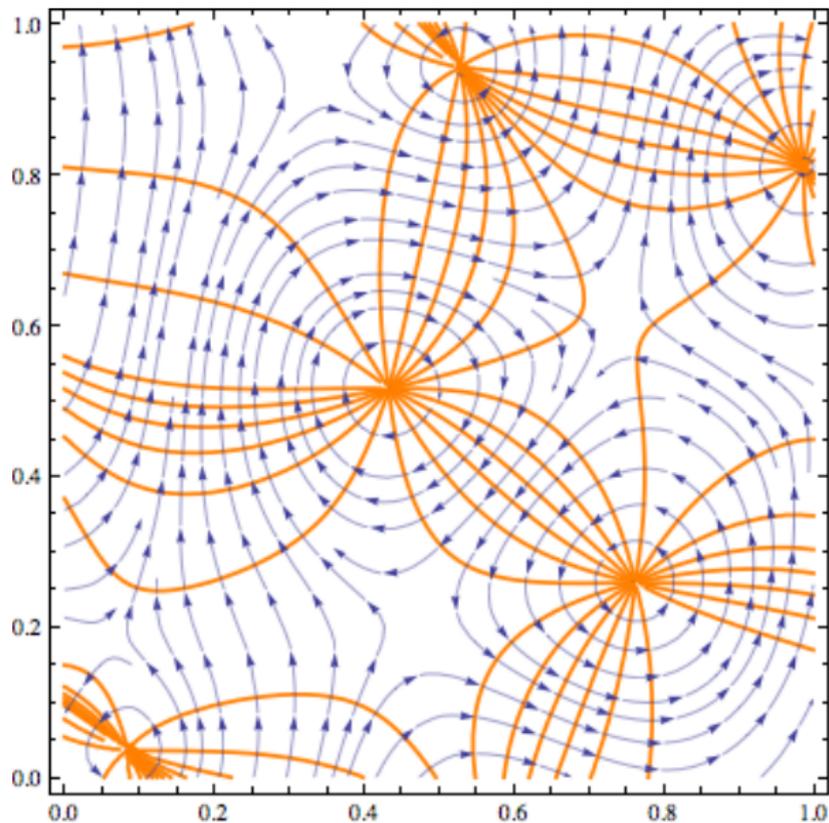There are two classes of curves:

- the integral curves of the phase field;
- the isophase curves called "wavefronts" in optics.

The field orthogonal to the isophase curves is the gradient field of $\varphi(a)$.

But $\varphi(a)$ is undeterminate at the singularities. In such cases physicists use the current field

$$\mathcal{J} = r^2 \nabla \varphi = X \nabla Y - Y \nabla X$$

## Superposition of plane waves

To get phase fields with a characteristic length, it is convenient to use superpositions of plane waves sharing the same wave number $k$: $Ae^{i\kappa.a}$

$A = Ee^{i\phi}$ complex amplitude

$\kappa = (\kappa_x, \kappa_y)$ wave vector

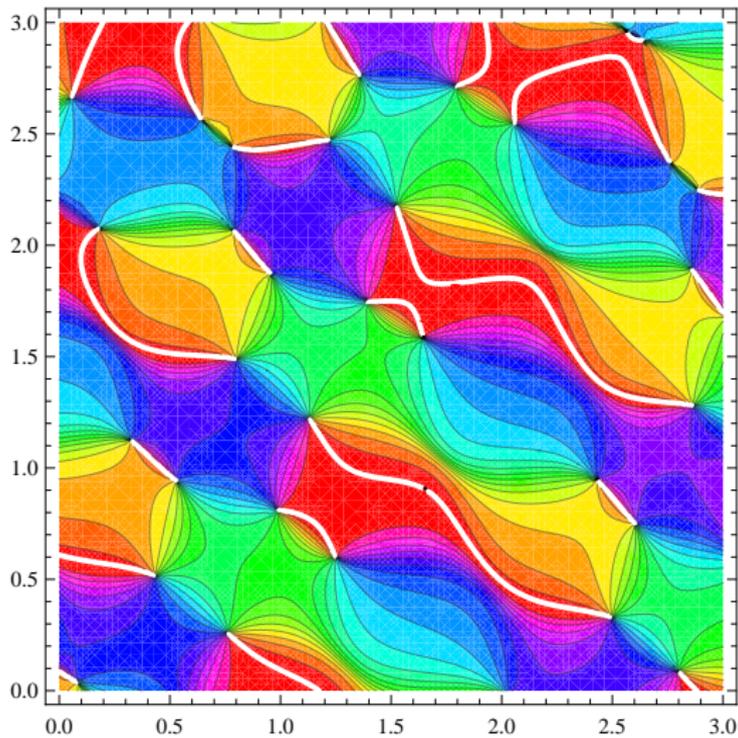$k = |\kappa|$ wave number

$\Lambda = \frac{2\pi}{k}$ wave length

They are solutions of the Helmholtz equation :

$$\Delta Z + k^2 Z = 0 .$$
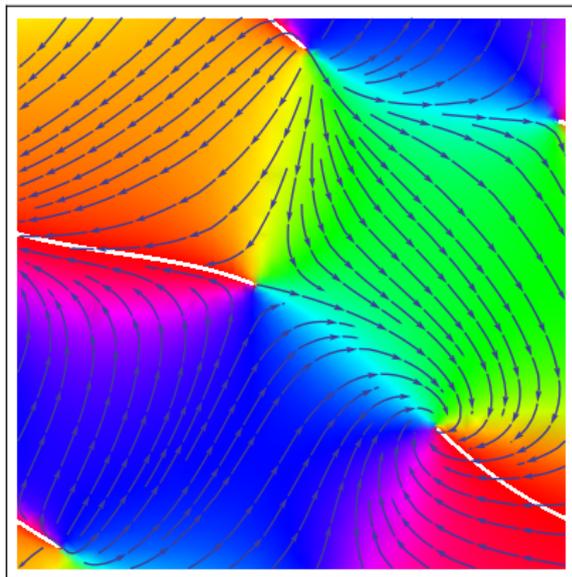
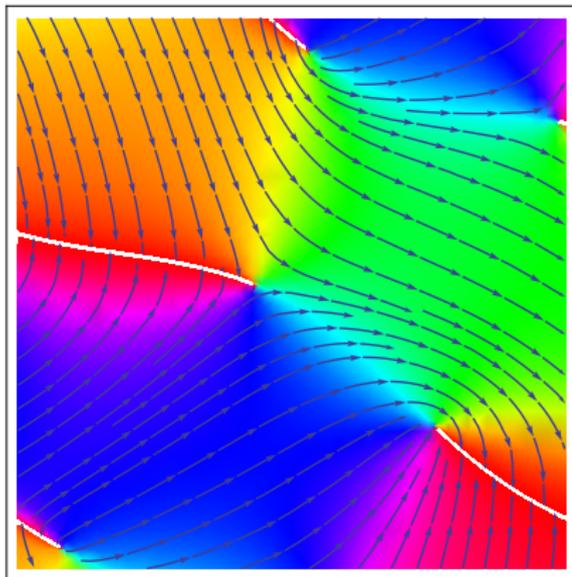Daniel Bennequin worked also on these models.

# Helmholtz pinwheels

Solutions have pinwheel-like isophase lines and provide very good models of empirical pinwheels..

When you look at underlying orientation integral lines, you get again end points and triple points.

But many experiments show that orientation selectivity does not vanish at the singular points $a_i$.

To take into account this key fact we can consider that pinwheels are local blow ups of points $a_i$ and look at the orientation field as the closure of a section $\sigma$ of $\pi : \mathbb{V} = \mathbb{R}^2 \times P \to \mathbb{R}^2$ defined over the open subset $\mathbb{R}^2 - \{a_i\}$. Over the $a_i$ the closure of $\sigma$ is the fiber $P_{a_i}$

At the limit, when all the points of the base plane $\mathbb{R}^2$ are blown up in parallel we get the fibration $\pi : \mathbb{V} = \mathbb{R}^2 \times P \to \mathbb{R}^2$.

So $\mathbb{V}$ can be considered as an idealized continuous approximation of the concrete $V1$. This model is now commonly used by neurophysiologists.

So, the idea is that pinwheels are local blow-ups of points $a$ whose exceptional fiber $P_a$ is compactified *à la* Kaluza-Klein, and projected in an "infinitesimal" neighborhood around $a$ in the base plane.
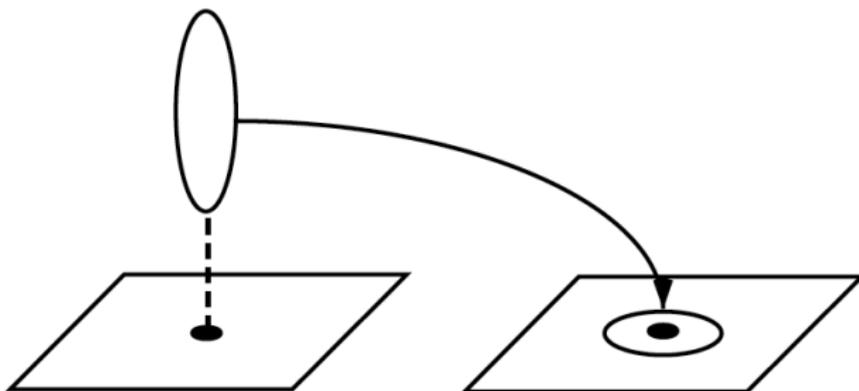
By the way, the *intuition* (not the mathematical concept) of a fiber bundle was explicit in Hubel with his concept of "engrafted" variables:

"What the cortex does is map not just two but many variables on its two-dimensional surface. It does so by selecting as the basic parameters the two variables that specify the visual field coordinates (...), and on this map it engrafts other variables, such as orientation..."

## Contact elements and 1-jets

Now, the elements of $\mathbb{V} = \mathbb{R}^2 \times P$ are *contact elements* that is numerical values of *1-jets of smooth plane curves*. They can be processed by neural "point processors".

So, we can give a first answer to our initial question:

QUESTION:
How the visual brain can be a neural geometric engine?

ANSWER:
Low dimensional *jet spaces* are neurally implemented and jet spaces are naturally endowed with integrability conditions.

Jets are "prolongations" in the sense of Cartan.

The visual brain is a "Lie-Cartan" geometric engine.

But we must go much further. Jets have to be *integrated* and this requires a supplementary structure.

The very key point is that this supplementary geometric structure on $\mathbb{V}$ is *implemented* in a specific class of neural connections.

Indeed, cortical neurons are connected by "horizontal" cortico-cortical connections *inside the layer itself*.

This supplementary connectivity is extremely *specific* and provides the *second part* of the functional architecture (the first part is provided the retino-geniculo-cortical "vertical" connections modeled by the fibration $\mathbb{V} = \mathbb{R}^2 \times P$).

The "vertical" retinotopic structure is not sufficient. To implement a *global* coherence, the visual system must be able to *compare* two retinotopically neighboring fibers $P_a$ and $P_b$ over two different base points $a$ and $b$.
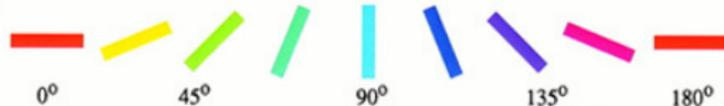
It is this problem of *parallel transport* which is solved by the *long-range* excitatory "*horizontal*" cortico-cortical connections.

# "Horizontal" cortico-cortical connections

Bosking's image shows the diffusion of a marker (biocytin) along horizontal connections (black marks). The injection site is upper-left in a green domain.
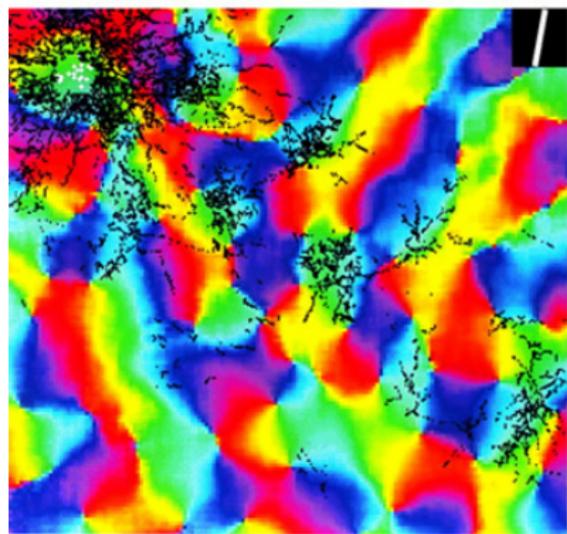
There are two main results:

1. the marked axons and synaptic buttons cluster in domains of the *same* color (same orientation), which means that horizontal connections implement neurally a *parallel transport*.

2. the global clustering along the upper-left bottom-right diagonal means that horizontal connections connect neurons with almost *parallel* and almost *aligned* orientations.

"The system of long-range horizontal connections can be summarized as preferentially linking neurons with co-oriented, co-axially aligned receptive fields." (W. Bosking)

This result is corroborated by experiments in psychophysics about what is called the *association field* and curve integration (David Field, Anthony Hayes and Robert Hess, Jean Loranceau).
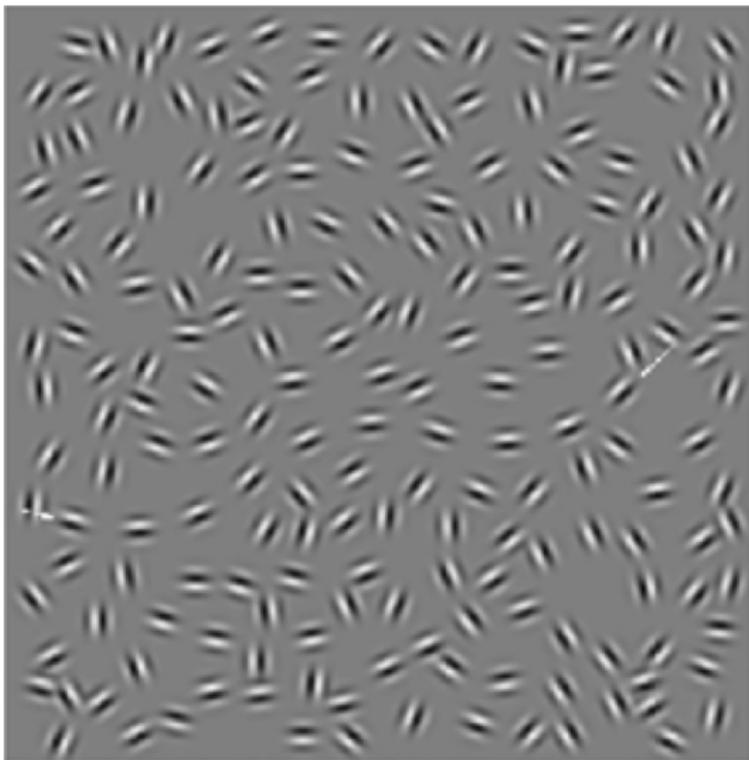
It is the neural origin of the the concept of a line.

The experiments concern the *pop-out* (the *perceptive saliency*) of almost aligned Gabor patches.

Here is an example. In a background of random patches, you insert a set of almost aligned patches and a global curve emerge.

A set of contact elements $c_i = (a_i, p_i)$ is perceived as a *global* curve (what is called a *binding*) if the orientations $p_i$ are *tangent* to a regular curve $\gamma$ interpolating as straightly as possible between the positions $a_i$.

These "joint constraints of position and orientation" (Field *et al.*) correspond to the horizontal connections.

horizontal connections, parallel transport, coaxiality, binding, propagation of coherent activity, synchronization, global pop out, saliency

$$\Longleftrightarrow$$

integration of contact elements

These experimental results show that a skew curve

$$\Gamma = v(s) = (a(s), p(s)) = (x(s), y(s), p(s))$$

in $\mathbb{V}$ is *perceived* as a globally coherent curve (via binding and pop-out) in the base plane $\mathbb{R}^2$

- iff $p(s)$ is the tangent $p = dy/dx$ to the curve $a(s)$,
- iff it is the *Legendrian lift* of its projection $\gamma$,
- iff it is an *integral curve* of the *contact structure* $\mathcal{K} = ker(\omega)$ of $\mathbb{V}$, where $\omega$ is the 1-form

$$\omega = dy - pdx$$

The contact structure is the kernel distribution of $\omega$, i.e. the distribution of contact tangent planes $\mathcal{K} = ker(\omega)$.

$\mathcal{K}$ is maximally non integrable since the 3-form

$$\omega \wedge d\omega = (-pdx + dy) \wedge dx \wedge dp = -dx \wedge dy \wedge dp \ .$$

is a volume form, which is the opposite of the Frobenius integrability condition $\omega \wedge d\omega = 0$.

So, even if there exists a lot of integral *curves* of $\mathcal{K}$ (Legendrian lifts), there exists no integral *surface*.

For curves, to work in $V1$ is to work with Legendrian curves.

The functional architecture is represented by the 1-form $\omega = dy - pdx$ and the associated contact structure.

The horizontal cortico-cortical connections mean geometrically that

the contact structure $\mathcal{K}$ of the space of 1-jets $\mathbb{V}$ is neurally implemented.

This explains how "point processors" as neurons can do differential geometry if they are connected by a suitable functional architecture.

We can go further.

The contact structure $\mathcal{K}$ is left-invariant for a group law making $\mathbb{V}$ isomorphic to the *polarized Heisenberg group*.

$$(x, y, p).(x', y', p') = (x + x', y + y' + px', p + p') .$$

Its Lie algebra is generated by the basis of left-invariant fields $X_1 = \frac{\partial}{\partial x} + p \frac{\partial}{\partial y} = (1, p, 0)$ and $X_2 = \frac{\partial}{\partial p} = (0, 0, 1)$ with $[X_1, X_2] = (0, -1, 0) = -\frac{\partial}{\partial y} = -X_3$ (the other brackets $= 0$).

The fact that the basis $\{X_1, X_2\}$ of the distribution $\mathcal{K}$ of contact planes is *bracket generating* (*i.e.* Lie-generates the whole tangent bundle $T^*\mathbb{V}$) is called the *Hörmander condition*.

Moreover $\mathbb{V}$ is *nilpotent* of step 2, which means that all brackets of the form $[t, [u, v]]$ vanish. It is an example of *Carnot group*.

# The contact structure of $SE(2)$

The Euclidean group $SE(2) = \mathbb{R}^2 \rtimes SO(2)$ of direct isometries of the plane acts naturally on $\mathbb{V}$ and it is therefore better to work in the associated principal bundle, as proposed by Giovanna Citti and Alessandro Sarti.

In that case, the contact form is

$$\omega_S = -\sin(\theta)\, dx + \cos(\theta)\, dy$$

that is $\cos(\theta)(dy - p dx) = \cos(\theta)\, \omega$.

The contact planes are spanned by the tangent vectors
$X_1 = \cos(\theta) \frac{\partial}{\partial x} + \sin(\theta) \frac{\partial}{\partial y}$ and $X_2 = \frac{\partial}{\partial \theta}$ with Lie bracket
$[X_1, X_2] = \sin(\theta) \frac{\partial}{\partial x} - \cos(\theta) \frac{\partial}{\partial y} = -X_3$.

Contrary to the polarised Heisenberg case, the $X_j$ constitute an Euclidean orthonormal basis and are therefore more natural.

The distribution $\mathcal{K}$ of contact planes is still bracket generating (Hörmander condition).

But $SE(2)$ is no longer nilpotent. In fact, the Carnot group $\mathbb{V}$ is in some sense "*tangent*" to $SE(2)$. It is called the "*tangent cone*" of $SE(2)$ or its "*nilpotentisation*".

That the functional architecture implements a contact structure explains some strange perceptive phenomena of very long range completion of images.

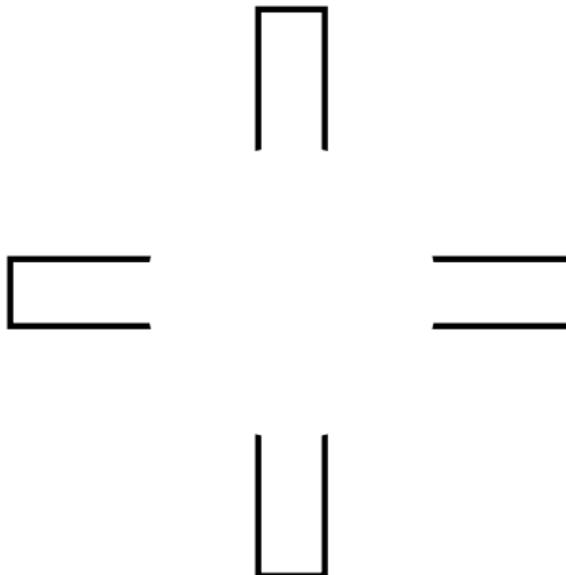Consider for example this Kanizsa square.
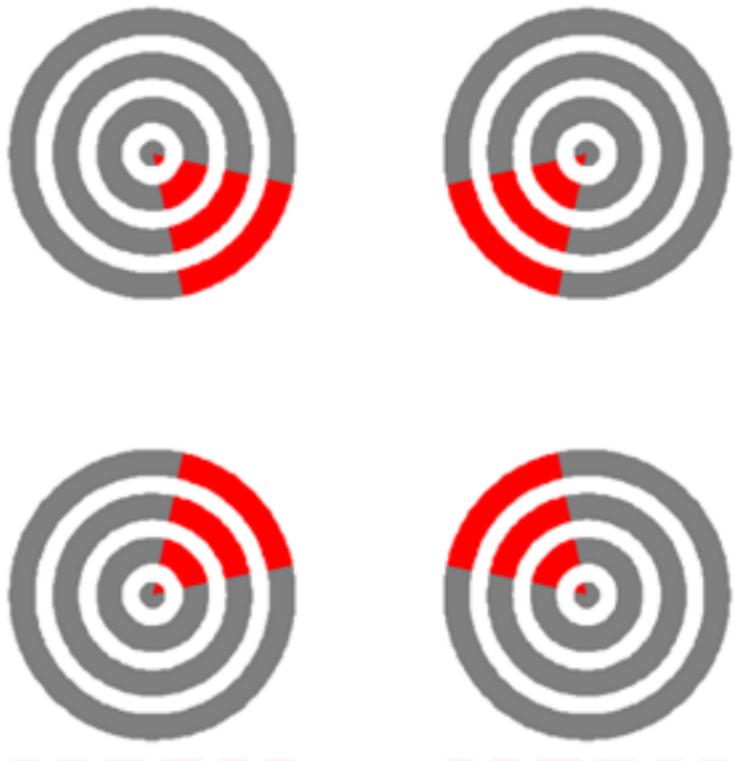
The red pacmen induce *very long-range curved illusory contours* (what is called modal completion).

Moreover, these contours act as *boundaries for a diffusion of color* inside the square (what is called the "neon" or "watercolor effect").

Consider also the Koffka cross:

The end points activate locally (area V2 is necessary) the orthogonal orientations.

These very sparse local activation induce a *very long-range global modal subjective contour*.

Moreover, subjects perceive alternately circles and squares, which means that there exists a *competition* between two completion strategies:

- *circle*: illusory contours with a maximal diffusion of curvature,
- *square*: piecewise linear illusory contours (curvature $= 0$) with corners (singularities of curvature).

This bistability is even more striking in the Ehrenstein illusion:
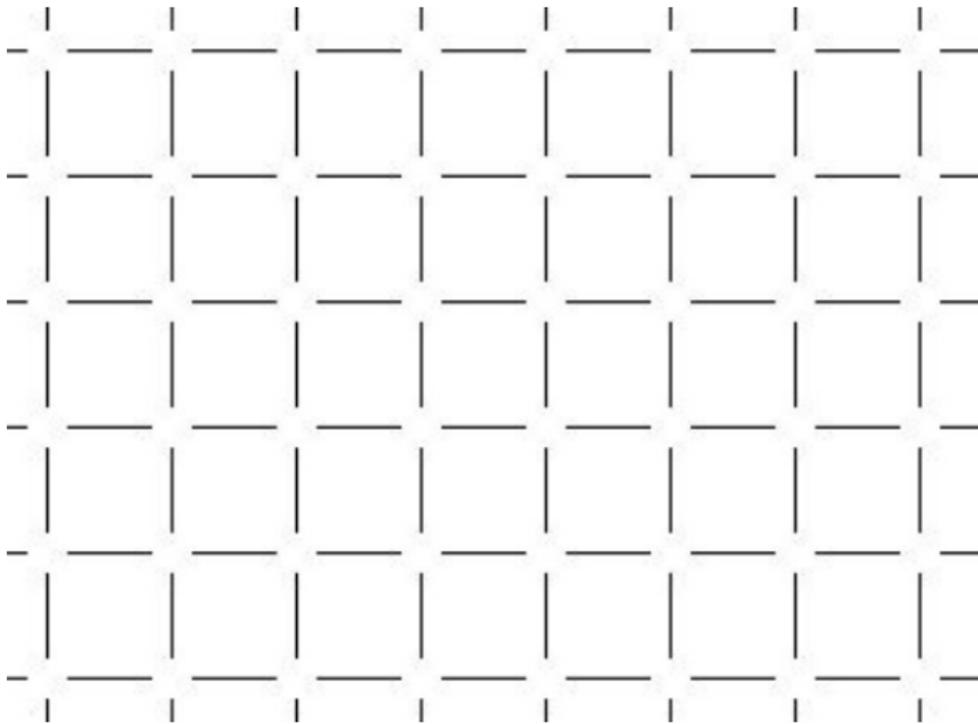
## Variational models

To explain these spectacular completion phenomena, *variational models* have been introduced since the late 70s.

They were models minimizing an energy along curves $\gamma$ *in the base plane* $\mathbb{R}^2$.

The best known is the *elastica* model proposed in 1992 by David Mumford. The energy to minimize is:

$$E = \int_\gamma (\alpha \kappa^2 + \beta) ds$$

But in what concerns neural models (and not only 2D image processing) it is natural to work in $V1$, that is with the *contact structure* and the *Legendrian curves*.

Hence the natural idea of introducing natural *sub-Riemannian* metrics on $\mathbb{V}$ and $SE(2)$ and look at *geodesic models* for curve completion and illusory modal contours.

The natural sub-Riemannian metrics on the distribution of contact planes $\mathcal{K}$ are the left invariant ones left translating the Euclidean metric of the contact plane at the origin.

The subRiemannian geometry of groups such as $\mathbb{V}$ and $SE(2)$ is rather complex. For the Heisenberg group the problem was solved in the 1980s by Richard Beals, Bernard Gaveau and Peter Greiner. They claimed:

> "The results indicate how complicated a control problem can become, even in the simplest situation."

A lot of geometers and analysts worked in this field. My main references have been Misha Gromov, Andrei Agrachev, John Mitchell, Richard Montgomery, Robert Strichartz, Anatoly Vershik, Jean-Pierre Pansu, Jean-Michel Bismut, André Bellaïche, Jean-Jacques Risler.

The structure of geodesics of $\mathbb{V}$ implies that the sub-Riemannian sphere $S$ (globally minimizing geodesics of SR length 1) and the wave front $W$ (only locally minimizing geodesics of SR length 1) are rather strange. One can compute them explicitly .

It is a control problem.

Due to the "Pontryagin maximum principle", geodesics are the projections on $\mathbb{V}$ of the Hamiltonian on the cotangent space

$$H\left(x, y, p, \xi^*, \eta^*, \pi^*\right) = \frac{1}{2}\left[(\xi^* + p\eta^*)^2 + \pi^{*2}\right] \ .$$

# The sub-Riemannian wavefront

The sphere $S$ and the wave front $W$ (with radius $\sqrt{2}$) are given by the equations

$$
\begin{cases}
x_1 = \frac{|\sin(\varphi)|}{\varphi} \cos(\theta) \\
p_1 = \frac{|\sin(\varphi)|}{\varphi} \sin(\theta) \\
y_1 = \frac{1}{2} x_1 p_1 + \frac{\varphi - \sin(\varphi)\cos(\varphi)}{4\varphi^2} \\
\quad = \frac{1}{2} \frac{\sin^2(\varphi)}{\varphi^2} \cos(\theta) \sin(\theta) + \frac{\varphi - \cos(\varphi)\sin(\varphi)}{4\varphi^2} \\
\quad = \frac{\varphi + 2\sin^2(\varphi)\cos(\theta)\sin(\theta) - \cos(\varphi)\sin(\varphi)}{4\varphi^2}
\end{cases}
$$

In the figure, the external surface is the sub-Riemannian sphere $S$. It has a saddle form with singularities at the intersections with the $y$-axis. The internal part is $W - S$. It presents smaller and smaller circles of cusp singularities which converge to 0. Such a complex behaviour is impossible in Riemannian geometry.
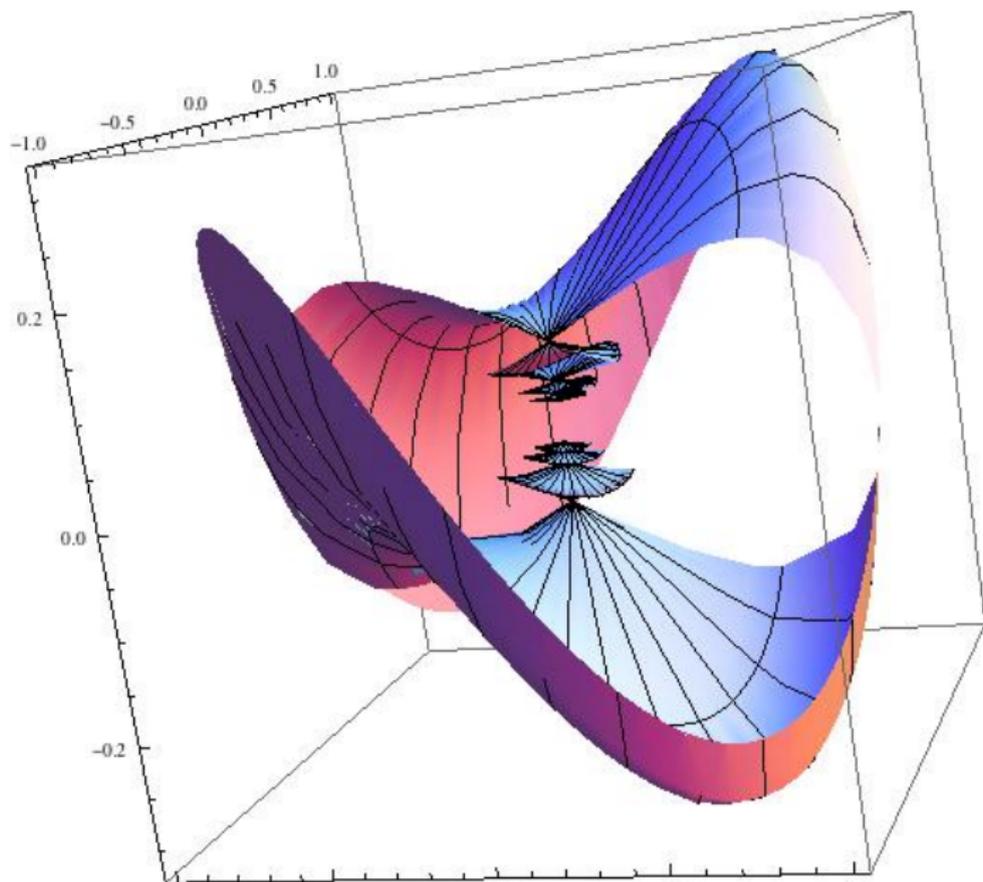
For $SE(2)$, the sub-Riemannian geometry is much more complex
and has been studied by the group of Andrei Agrachev, in
particular Jean-Paul Gauthier, Ugo Boscain and Yuri Sachkov.

Giovanna Citti, Alessandro Sarti and Remco Duits studied also the
problem.

So virtual visual contours are integrated via sub-Riemannian geodesics implemented in horizontal connections.

For the completion of corrupted images (inpainting) it is therefore natural to use *diffusion* along the horizontal connections, that is sub-Riemannian Laplacian, SR heat kernel, etc.

The numerical results are quite striking.

The following picture due to Jean-Paul Gauthier shows how a highly corrupted initial image can be very well restored using sub-Riemannian diffusion.
Top-left: initial image, top-right: highly corrupted image, bottom: restored image.
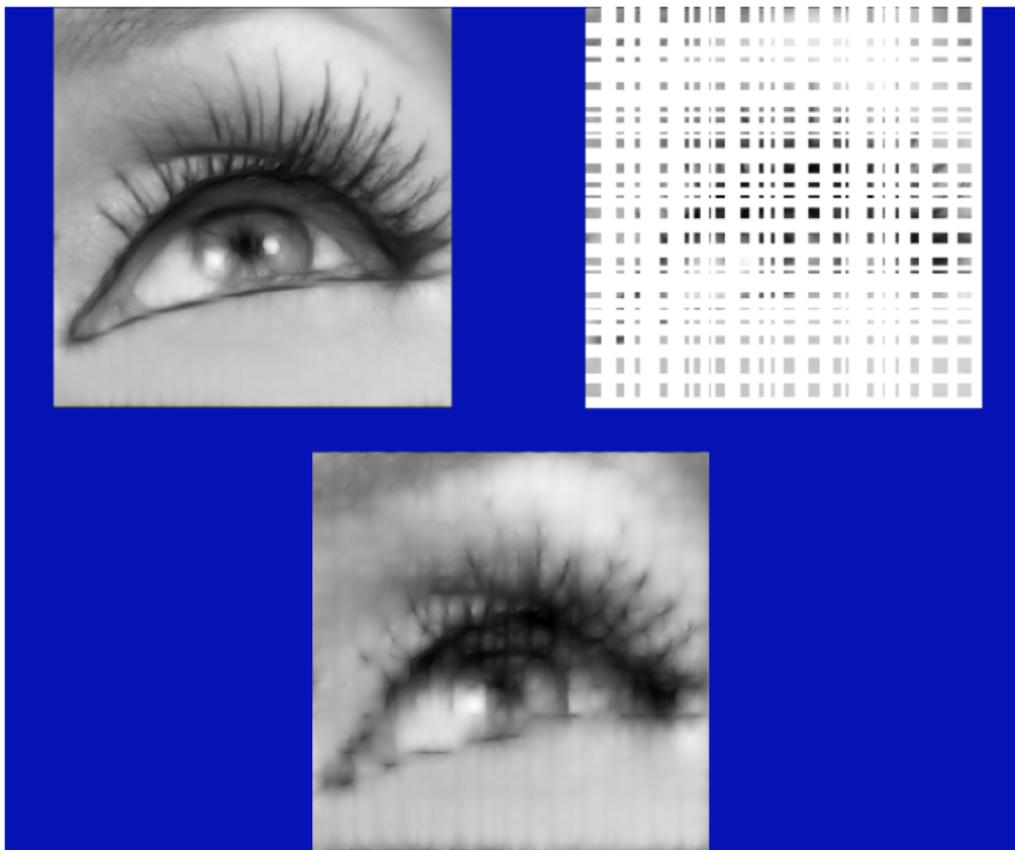
Happy Birthday, Bernard

and a warm thanks for your support to cognitive neurosciences